BODC Project Database Structure

PARENT HOME

Introduction

A relational database is made up from tables. Each table contains one or more fields. Some of these fields, termed key fields, are contained in more than one table and provide the mechanism for linking tables, and hence the data they contain, together. If a key field occurs once in each of two tables, then a one to one relationship between the tables is established. If the key field occurs once in one table and many times in the other, then a one to many relationship exists between the tables.

The BODC Community Research Project databases are relational databases that have been built for the specific task of storing project data in a way where it may easily be found when required. The end result is an effective, if not elegant, design.

Simple structures exist that match the data. These structures are extended in response to data sets supplied. However, as the databases have developed, certain patterns have been recognised in the data. In response to this, fully normalised structures (i.e. the type conventional database designers desire) have been developed. The advantage of these normalised structures is that providing the rules of their underlying data model are obeyed by the data, their scope may be expanded with no maintenance overheads. Their disadvantage is that significantly more work is required getting data in and getting data out.

We therefore have a situation in practice where simple and fully normalised structures exist side by side. Any spare resource is directed towards converting the simple structures into normalised structures, providing a clear advantage can be seen in doing the work. The structure of the database is therefore dynamic but it is supported by 'soft' documentation that can evolve in parallel.

Table Types

The database may be considered as containing five types of table. The database is built on an event-based data model. In other words, something has to happen to generate the data stored. The primary information in the database therefore has to describe what these events were, where they happened and when they happened.

This information is stored in the database primary index tables. In the following table definitions, the hot links to these tables are coloured red.

The data model assumes that the events are related to the data they generate by one to many relationships. These relationships are implemented in the database by one or more secondary index tables. These tables also provide storage for metadata that are specific to a single type of event. The hot links to these tables in the table definition index are coloured green.

The third type of table is the fully normalised data table. These may be regarded as stable, long term entities within the database. Because the structures are normalised, it is not possible to obtain the sort of cross-tabulated output most users require using simple

SQL queries. Consequently, data access tools are provided by BODC. The hot links to these tables in the table definition index are coloured blue.

The normalised data tables are supported by a series of code tables, such as the parameter dictionary, that together may be termed the dictionary tables. The hot links to these tables in the table definition index are coloured 'dark yellow'.

Finally, there are the simple-structured data tables. These may be effectively interrogated by simple SQL queries. However, they should be regarded as transient entities that may disappear from subsequent database releases. Obviously, if they do disappear, the data they contained will have been transferred to a fully normalised structure within the database. The hot links to these tables in the table definition index are coloured magenta (pale purple).

The Parameter Dictionary

The parameter dictionary is an essential feature of the normalised data storage tables. The identification of parameters is based on 8-byte codes. These have been designed using a hierarchical model. The first four bytes may be considered as the 'parent'. This provides information on the parameter at a low level of detail. This parent has one or more 'child' 8-byte codes. These subdivide the parent into more detailed information.

This relationship is exploited in different ways. For example, chemical parameters have a parent field identifying the basic parameter with the children identifying different measurement protocols. Thus 'CPHL' is chlorophyll-a, but CPHLHPP1 is chlorophyll-a measured by reverse-phase HPLC on an acetone extract from a GF/F filter. For biological species codes the parent specifies the genus and the children the species.

There are a large number (thousands) of parameters coded in the database. Finding out what a given code means is straightforward (!). A query matching on field CPMUSG of table ZUSG will provide the answer. However, specifying the parameter code for data retrievals requires some thought.

The secret lies in the use of wild cards which any database management system can incorporate into query searches. The recommended technique is to use table ZUPM to identify the parent code of the parameter you require. A wild card may then be set up to include as many of the child codes as required. One word of warning. Always check the meanings of all codes covered by a wild card as there are traps for the unwary. For example, the wild card CORG% covers both CORGCAP1 ("POC") and CORGCOD1 (DOC) which should not be merged into a single data set!

Documentation Structure

This document contains two main sections. These are:

Table Definitions: A description of the fields contained in each table of the database relevant to the ARABESQUE project.
Linkage Definitions: Documentation that describes how the tables of the database are linked together through their key fields.

TIP If you are looking for a particular type of data and don't know which tables you require, looking through the linkage definitions will provide a quick and easy way of finding out what you need to know.

Database Table Definitions

This section provides a field level description of all the user-accessible tables in the OMEX and ODB databases which are relevant to the ARABESQUE project. A complete description of ALL the database tables is given in the version of this document found on the OMEX I CD-ROM..

Table BINCTD

This table contains the CTD profile data, averaged into either 1 db (casts shallower than 100 db) or 2 db bins.

Table BOTDATA

This table contains analytical data on water and air samples. A very wide range of parameters is stored here.

Table BOTTLE

This table provides an inventory of the water and air samples, collected by a variety of methods, held in the database. Note that the name BOTTLE is more of a historical relic than a description of current usage. Vital information, particularly the depth or height from which the sample was taken, is held in the table.

Table CRSINDX

This table allows the user to identify which events were associated with which project by linking together the cruise fields from table EVENT. Metadata storage fields for each cruise are also provided.

Table CTDCAL

This table contains the calibration coefficients applied to the CTD profiles by the BODC processing system.

Table CTDINDX

This table provides an inventory of the CTD casts held in the database together with storage for CTD-specific information.

Table CTDTYP

This is a code table that supports table CTDINDX by defining the mnemonics used to identify CTD instrument types.

Table EVENT

An event is defined as any activity that results in the collection of data that are stored in the database. Table EVENT contains information on what the event was and where and when it occurred. It could therefore be considered as the most important table in the database and should certainly be involved at the start of any search for data.

Table EVENT_COMM

This is an extension of the EVENT table that carries a plain language comment field. This is only separated to make the EVENT table less cumbersome to list.

Table G_CODE

G_CODE is a simple code table that defines the gear codes used in table EVENT.

Table OXYDAT

This table contains TCO_2 and oxygen uptake (gross and net production) data from long period (generally 24 hour) on-deck and in-situ incubations, including size-fractionated data.

Table OXYHDR

This table provides an inventory of the TCO_2/ oxygen uptake (gross and net production) uptake experiments held in the database. Vital metadata fields are included. The table also provides storage for column integrated data.

Table ORGCODE

Table ORGCODE is a code table defining the data originator codes used in the database.

Table PRINDX

This table provides an inventory of light profiles collected by profiling radiometers and the link between the EVENT and PRPROF tables. Storage is also provided for light profile metadata parameters.

Table PRPROF

This table is used to store light profile data as measured by various types of profiling radiometers.

Table ZUCT

This table is part of the parameter dictionary. Its function is to group the 4-byte 'parent' parameter codes into categories to enable them to be interrogated more easily.

Table ZUNT

This table is part of the parameter dictionary. It is a code table that defines the codes used to represent parameter units.

Table ZUPM

This table is part of the parameter dictionary. It contains the definitions of the 4-byte 'parent' parameter codes (i.e. the first four bytes of the parameter code).

Table ZUSG

This is the main table of the parameter dictionary, containing definitions of the full 8-byte parameter codes.

Table BINCTD
Field Definitions

BEN	NUMBER(6)	BODC event number.
PRESS	NUMBER(5,1)	Pressure (db).
TEMP	NUMBER(5,3)	Temperature (°C).
SALIN	NUMBER(5,3)	Practical salinity (PSU).
SIGMA	NUMBER(5,3)	Potential density anomaly (kg/m 3).
O2	NUMBER(4,1)	Dissolved oxygen (µM).
O2SAT	NUMBER(4,1)	Oxygen saturation (%).
CPHYL	NUMBER(4,2)	Chlorophyll (mg/m 3).
ATTEN	NUMBER(5,3)	Optical attenuance (per m).
DWIR	NUMBER(5,1)	Downwelling irradiance (µE/m 2 /s).
UWIR	NUMBER(4,1)	Upwelling irradiance (µE/m 2 /s).
POTEMP	NUMBER(5,3)	Potential temperature (°C).
POAT	NUMBER(5,3)	Potential attenuance (per m).

Notes
The pressure value signifies the midpoint of the bin. Thus, a pressure of 1.0 db signifies a bin extending from 0 db to 2 db (assuming that the cast was deeper than 100 db).

The density parameter computed is the potential density anomaly calculated at 0 db and is numerically equivalent to the parameter known as sigma-theta (computed by substituting potential temperature into the UNESCO SVAN function).

Oxygen saturation has been computed using the algorithm of Benson and Krause (1984).

UNIX users should note that CTD data are only loaded into BINCTD once BODC has full confidence in the CTD calibrations. Our normal practice is to leave this operation until near the end of the project to allow the maximum time for feedback from the user community. Prior to loading into BINCTD, the data are held uncalibrated in holding tables that are inaccessible to users through SQL. If these data are required, the CTDLIDST utility should be used which will retrieve the data and dynamically apply any calibrations required. For CD-ROM database releases, all CTD data have been incorporated into BINCTD.

Reference
Benson, B.B., Krause D. (1984). The concentration and isotopic fractionation of oxygen dissolved in fresh water and sea water in equilibrium with the atmosphere.
Limnol.Oceanogr., 29, 620-632.

Table BOTDATA
Field Definitions

IBTTLE	NUMBER(6)	BODC bottle/sample identifier.
CPCODE	CHAR(8)	Parameter code.
FPVAL	NUMBER	Parameter value.
CPFLAG	CHAR(1)	Parameter quality control flag.
IORGRF	NUMBER(6)	Originator's reference.
IDOCRF	NUMBER(8)	Document reference.
CILOAD	CHAR(6)	Record creation date (yymmdd).
TSGMOD	DATE	Last modification time stamp.

Notes
The primary key is formed from the three fields, IBTTLE, CPCODE and IORGRF. In other words, the table contains one row for each parameter measurement on each water or air sample by a given data originator. The parameter code consists of 8 bytes which describe the parameter measured in some detail. The parameter code definitions are stored in the parameter dictionary (see the table names starting with 'Z').

The parameter flag field serves two purposes. First, it identifies parameter values identified as problems during quality control procedures. Different codes are used to differentiate between originator, BODC and user quality control. Secondly, it is used to identify samples where the measured parameter was either below detection limit or saturated the measuring apparatus. In these cases the data values are set to the detection limit (zero if no detection limit was specified) or the saturation value respectively. If no flag value has been assigned (signifying good data), the CPFLAG field is blank.

The flag values which may be encountered are:

K Uncertain/suspect value (source of quality control unknown).
L Uncertain/suspect value (data originator's quality control).
M Uncertain/suspect value (BODC quality control).
O Uncertain/suspect value (user quality control)
T Nearest value to bottle firing depth
< Below detection limit.
> In excess of stated value.

The 'T' flag is only found on records created for water bottle samples from CTD profile data. It means that no data were found at the bottle firing pressure. Instead, the nearest data value has been taken, providing this was within 2 db of the required pressure.

The originator's reference field allows the suppliers of individual data values to be identified. The objective when allocating these linkages is to provide a point of contact for users of the data to approach when initiating collaboration that will endure beyond the end of a project. Consequently, linkages have been assigned at the PI level and do not necessarily specify the individual who actually did the analysis.

The capability to link data to its originator only came about when the normalised structure was implemented. Linkages have been retrospectively applied to the entire data holding during restructuring which was done using cruise reports and the collective memories of BODC staff and participating scientists. If we’ve got anything wrong, please don’t bear a grudge: just let us know and we’ll fix it. Likewise, anyone who feels aggrieved for any reason about these code allocations should discuss it with us so that any problems may be quickly rectified.

Codes are used to eliminate potential problems with misspellings and the like. The codes used are documented in the table ORGCODE.

Document references have not yet been implemented so the IDOCRF field is currently always null.

Table BOTTLE
Field Definitions

BEN	NUMBER(6)	BODC event number.
IBTTLE	NUMBER(6)	A unique identifier assigned by BODC to each sample.
MINP	NUMBER(5,1)	Minimum pressure for the sample (db).
MAXP	NUMBER(5,1)	Maximum pressure for the sample (db).
DEPTH	NUMBER(6,2)	Sampling depth (m).
BOTYP	CHAR(4)	Bottle/sample type identifier.
FLAG	CHAR(1)	Problem indicator flag.

Notes
Table BOTTLE was originally conceived for the management of water bottle data. However, as the BODC databases developed, it was realised that the table could be utilised for other data types. Data currently held include pumped air and water samples, stand-alone pump (SAP) samples, bucket samples and air bottle samples.

The most important function of this table is to implement the 'one-to-many' relationship that may exist between samples and events. The table contains one row per sampling depth (multiple samples at a single depth are considered as one).

Each record in EVENT can 'own' as many records as it likes in BOTTLE through the foreign key field BEN. Hence each EVENT can include many sampling depths.

The relationship between MINP, MAXP and DEPTH requires some explanation. MINP and MAXP only have relevance to bottles on a CTD rosette. In this case, bottle 'depths' are frequently logged as pressure ranges during CTD screening and loaded into BOTTLE. Subsequently, DEPTH (the distance from the surface to the midpoint of the bottle) is derived by applying a pressure calibration to MINP and MAXP, correcting for CTD frame geometry, and applying the standard conversion from pressure to depth. In order to allow for pressure calibration drift, the minimum value is constrained at 0.5 m. The fields MINP and MAXP provide a direct linkage between BOTTLE and the CTD data which is why they are retained. For other sample types, DEPTH is assigned a value from reports or logs and MINP and MAXP are left null. Note that air samples have negative depths to indicate height above sea level.

The BOTYP field specifies how the sample was collected. For water bottle data, this field identifies the type of water bottle used through the following codes.

GFnn GoFlo
LNnn Lever-action Niskin (external spring)
NInn Niskin
NOnn NIO bottle
NXnn NOEX bottle
TRnn Transparent (marine snow catcher)
PPnn Manually filled polypropylene bottle
KNnn Knudsen bottle
G300 Large volume (300 litre) radionuclide sampler

The 'nn' specifies the capacity of the water bottle in litres.

A number of other codes are used for other sampling methods:

BWS Benthic water sampler
SAP Stand-alone pump
PUMP Pumped sample (water or air pump)
BUCK Bucket on a rope
AIRB Air bottle

The FLAG field is used to indicate known problems. The coding convention used is:

B Filter burst (SAP samples)
L Contamination through leakage suspected
M No sample obtained
O Bottles fired in incorrect order

The 'O' flag requires a little more explanation. This is used to flag stations where there was obvious confusion from the sample data set about which bottle was fired at which depth. These problems have been resolved during data load, but the flag is included to remind users that there may be problems with data from that station obtained outside the database.

Table CRSINDX
Field Definitions

CRUISE	CHAR(8)	BODC cruise mnemonic,
PROJECT	CHAR(12)	Mnemonic of the project with which the cruise was associated.
PSO	CHAR(20)	Cruise chief scientist.
COUNTRY	CHAR(20)	Country responsible for organizing the cruise.
TBEGNS	DATE	Date the cruise sailed.
TENDS	DATE	Date the cruise docked.
LOCATION	CHAR(80)	Plain language description of the area studied.
COMM	CHAR(60)	Plain language comment field.

Notes
This table allows events associated with a particular project to be identified as well as providing limited background information on cruises.

Table CTDCAL

Field Definitions

BEN	NUMBER(6)	BODC event number.
FPCOR	NUMBER(5,2)	Pressure correction (db).
FCSLOP	NUMBER(7,4)	Chlorophyll calibration slope.
FCIRR	NUMBER(8,7)	Chlorophyll calibration irradiance term.
FCCEPT	NUMBER(7,4)	Chlorophyll calibration intercept.
FUSLOP	NUMBER(7,4)	Upwelling irradiance calibration slope.
FUCEPT	NUMBER(7,4)	Upwelling irradiance calibration intercept.
FDSLOP	NUMBER(7,4)	Downwelling irradiance calibration slope.
FDCEPT	NUMBER(7,4)	Downwelling irradiance calibration intercept.
FBASOF	NUMBER(5,2)	Distance between the CTD pressure sensor and the base of the water bottle (m).
FTOPOF	NUMBER(5,2)	Distance between the CTD pressure sensor and the top of the water bottle (m).
FTEMOF	NUMBER(5,2)	Distance between the CTD pressure sensor and the reversing thermometer (m).
FSSLOP	NUMBER(7,5)	Salinity calibration slope.
FSCEPT	NUMBER(7,5)	Salinity calibration intercept.
FTSLOP	NUMBER(7,5)	Temperature calibration slope.
FTCEPT	NUMBER(7,5)	Temperature calibration intercept.
FOSLOP	NUMBER(6,3)	Oxygen calibration slope.
FOCEPT	NUMBER(6,3)	Oxygen calibration intercept.
FSMSLOP	NUMBER(7,5)	Total suspended matter calibration slope.
FSMCEPT	NUMBER(7,5)	Total suspended matter calibration intercept.
FOMSLOP	NUMBER(7,5)	Organic suspended matter calibration slope.
FOMCEPT	NUMBER(7,5)	Organic suspended matter calibration intercept.
FIMSLOP	NUMBER(7,5)	Inorganic suspended matter calibration slope.
FIMCEPT	NUMBER(7,5)	Inorganic suspended matter calibration intercept.

Notes

This table contains one row per CTD cast and therefore allows each CTD to have a separate calibration. However, in most cases calibrations have been set up on a cruise by cruise basis.

In the case of UK WOCE and some of the OMEX data, the CTD data are supplied to BODC fully calibrated by the data originator. In such cases, CTDINDX records are set up with dummy values which have been set up to ensure the correct functioning of the BODC ctdlist software. Some transformation of the data is also necessary in these cases. For example, log transforms are applied to parameters which have an exponential calibration applied by ctdlist.

Each calibration and its method of determination is now discussed.

Rig Geometry

The fields FBASOF, FTOPOF and FTEMOF contain the information required to compute the true water bottle depth from the CTD pressure channel. FTEMOF is used when extracting calibration temperatures. The water bottle depth for a given CTD pressure reading (calibrated and converted to depth) is given by:

BOTTLE DEPTH = CTD DEPTH - (FBASOF + ((FTOPOF - FBASOF)/2.0))

This equation assumes bottle depth to be defined as the depth to the midpoint of the water bottle. The depth of the reversing thermometer is obtained by simply subtracting FTEMOF from the calibrated and converted CTD pressure reading.

The values used in these fields were obtained from actual measurements of the CTD rigs.

Pressure

The pressure correction, FPCOR, is a simple offset which is added to the uncalibrated CTD pressure. It is derived by consideration of data logged when the CTD was obviously out of the water.

Temperature

The temperature calibration has two components, FTSLOP and FTCEPT, which are applied to the uncalibrated temperature using the equation:

TCAL = TRAW * FTSLOP + FTCEPT

The temperature calibration is derived by comparison of the CTD temperature channel with calibrated digital reversing thermometer data for a specified cruise. A mean offset is computed after rejection of suspect reversing thermometer readings and stations where the reversing thermometers were fired on a temperature gradient.

In most cases, the accuracy of CTD resistance thermometers exceeds that of the digital reversing thermometers in common use. Consequently, the calibration coefficients are set to 1 and zero unless a problem is suspected with the CTD calibration.

Salinity

The salinity calibration is identical in form to the temperature calibration and has been derived in a similar manner using water bottle samples assayed on a bench salinometer.

Chlorophyll

The chlorophyll concentration (in mg chlorophyll a/m ^{3}) may be obtained from the fluorometer voltage using the following equation:

CHLOROPHYLL = EXP (VOLTS * FCSLOP + FCIRR * DWIR + FCCEPT)

The chlorophyll calibrations were set up by multiple regression of fluorometer voltages and downwelling irradiance (at the water bottle firing depths) against the log of the associated extracted chlorophyll measurements. The calibration was done on quality controlled data.

It should be noted that, on some cruises, the FCIRR term is zero because no downwelling irradiance data were available (for example the DI183 calibration was done on samples taken from pre-dawn casts).

Oxygen

The oxygen calibration is of the form:

OXCAL = OXRAW * FOSLOP + FOCEPT

and is derived by regression of CTD channel values at the bottle firing depths against Winkler titration results. On cruises where underway oxygen data were available, the surface CTD values (averaged over the top 3m after screening) are compared with the calibrated underway oxygen values to provide additional calibration data.

Suspended Load

The suspended matter calibrations are of the form:

SUSPENDED LOAD = (ATTENUANCE - INTERCEPT) / SLOPE

The calibrations have been obtained by regression of beam attenuance against gravimetric determinations of suspended load. Organic sediment load was determined by loss on ignition of the gravimetric samples.

It should be noted that suspended matter calibrations on deep water transmissometer data are extremely rare as huge quantities of water need to be filtered to obtain the necessary gravimetric data.

Irradiance

The raw irradiance (upwelling and downwelling) data are held as voltages. These are calibrated using the equation:

IRRADIANCE = EXP (VOLTS*SLOPE + INTERCEPT)

This returns calibrated values in units of $\mu W/cm^2$

For the PML 2-pi PAR meters currently used on CTDs on the NERC ships, an empirical calibration factor (0.0375) has been determined to convert these data into $\mu E/m^2/s$. The calibration is only valid over the range -1.5V to +1V.

Table CTDINDX

Field Definitions

BEN	NUMBER(6)	BODC event number.
TBEGNC	DATE	Date/time of the start of the downcast.
TENDC	DATE	Date/time of the end of the downcast.
MAXP	NUMBER(5,1)	Maximum pressure in the downcast (db).
FMAXP	CHAR(1)	Set to 'C' if the pressure calibration held in table CTDCAL has been applied to MAXP. Otherwise left null.
EXTCO	NUMBER(5,3)	Downwelling irradiance extinction coefficient.
MLD	NUMBER(3,1)	Mixed layer depth (m).
EZD	NUMBER(4,1)	Depth to the base of the euphotic zone (m).
TYPE	CHAR(3)	Type code of the CTD used (i.e. NB3 for Neil Brown Mk. III).

The downcast start and end times have been derived from the CTD data time channel and may be used to regenerate that channel if required.

The fields EXTCO, MLD and EZD were set up for the BOFS programme. In practice, it has been found that providing universally acceptable algorithms for their computation is an impossible task. Consequently, current practice is to leave the fields null unless agreed values are provided by the scientific community.

Table CTDTYP
Field Definitions

TYPE	CHAR(3)	CTD type code mnemonic.
DESCR	CHAR(30)	Plain language definition of the mnemonic.

Table EVENT
Field Definitions

BEN	NUMBER(6)	BODC event number. A unique numerical identifier assigned each event.
OID	CHAR(12)	What the event was known as during the cruise(originator's identifier)
GCODE	CHAR(8)	Code used to specify the gear pertaining to the event.
TBEGNS	DATE	Event start date/time (GMT).
TENDS	DATE	Event end date/time (GMT).
LAT	NUMBER(7,5)	Average latitude for event (°+ve North).
LON	NUMBER(7,5)	Average longitude for deployment (°+ve East).
VARLAT	NUMBER(7,5)	Maximum deviation of latitude from mean during station.
VARLON	NUMBER(7,5)	Maximum deviation of longitude from mean during station.
WDEPTH	NUMBER(5,1)	Average bathymetric depth for the event (m).
LATS	NUMBER(7,5)	Latitude at time TBEGNS (°+ve North).
LONS	NUMBER(7,5)	Longitude at time TBEGNS (°+ve East).
LATE	NUMBER(7,5)	Latitude at time TENDS (°+ve North).
LONE	NUMBER(7,5)	Longitude at time TENDS (°+ve East).
CRUISE	CHAR(8)	Cruise mnemonic.
SITE	CHAR(12)	Fixed station name.

Notes

This table has been built from the best available information from cruise reports, log sheets and information accompanying data. Automatically logged navigation has been used to match times and positions wherever possible.

There are two types of event, point events and traverse events. Point events may be considered as those events that effectively happen at a fixed position. Their positions are specified by the fields LAT, LON, VARLAT and VARLON with the other four position fields left null. Traverse events, such as tows and trawls, involve the ship steaming a significant distance. In this case, the start and end positions are stored in LATS, LONS, LATE and LONE. Note that some point events have data entered into the point event position fields to allow them to be handled as very low resolution points as required. Water depths are only included for point events.

Wherever possible, the fields LAT and LON are derived by averaging the data from the ship's navigation log over the event duration. VARLAT and VARLON are the maximum deviation of the data set from the mean. If VARLAT and VARLON are null then the data in LAT and LON have been take from logs or reports.

Obviously, the average of the ship's positions are not used for moorings. If VARLAT and VARLON are set then the information has been derived from the difference of the recorded positions on deployment and recovery.

The BODC event number (BEN) is a concept introduced to overcome the problem that it is impossible to guarantee that the identifiers assigned during the cruise will be unique within database incorporating many cruises. It is a very important field because it is used within the database as a 'primary key' which by definition must be unique. Data elsewhere in the database, resulting from a specified event, will either be labelled directly, or via a linkage record to its BEN.

OID, the originator's identifier, is the label that was assigned to the event during the cruise. For example, for Discovery cruises, it is based on the 'Discovery number' such as 11869#1. In a few cases, usually non-toxic samples or XBT drops, no identifier was assigned during the cruise and suitable naming schemes have been devised by BODC.

Event start and end times have been specified to bracket the event. Thus, for a CTD cast, the time span is from the instrument leaving the deck until its return. Some events are regarded as instantaneous, for example non-toxic samples. In these cases, the end times are set null. Wherever possible, cores are regarded as instantaneous events at the time when the corer reached the bottom.

The gear codes are mnemonics used to describe the data collection activity or the equipment used. The codes have been chosen to convey as much meaning as possible, but a plain language description of each code is provided in table G_CODE.

The cruise identifiers are made up from a ship code concatenated with the cruise identifier. For example, 'DI' is used for Discovery and 'CD' for Charles Darwin and a typical cruise would be labelled DI182 or CD46.

Table EVENT_COMM
Field Definitions

BEN	NUMBER(6)	The BODC event number.
COMM	CHAR(100)	Plain language comment field.

Notes
This table provides a mechanism for labelling EVENT records without encumbering EVENT listings with a large text field.

Table G_CODE
Field Definitions

GCODE	CHAR(8)	Standardised gear code.
DESCR	CHAR(60)	Plain language description of the gear described by GCODE.

Table OXYDAT
Field Definitions

EXPREF	CHAR(6)	BODC experiment reference.
IBTTLE	NUMBER(6)	BODC bottle reference number.
DEPTH	NUMBER(4,1)	Depth (or depth equivalent) at which sample was incubated.
GOX	NUMBER(4,2)	Oxygen gross production (µM/day).
SGOX	NUMBER(4,2)	Standard deviation of gross oxygen production (µM/day).
NOX	NUMBER(4,2)	Oxygen nett production (µM/day).
SNOX	NUMBER(4,2)	Standard deviation of nett oxygen production (µM/day).
GTCO2	NUMBER(4,2)	TCO_2 gross production (µM/day).
SGTCO2	NUMBER(4,2)	Standard error of TCO_2 gross production (µM/day).
NTCO2	NUMBER(4,2)	TCO_2 nett production (µM/day).
SNTCO2	NUMBER(4,2)	Standard error of TCO_2 nett production (µM/day).
PLIGHT	NUMBER(3)	Fraction of available light illuminating the sample (%)
SRTCO2	NUMBER(4,2)	Standard error of TCO_2 respiration (µM/day).
GPCO2	NUMBER(4,2)	pCO_2 gross production (µM C/day).
SGPCO2	NUMBER(4,2)	Standard error of pCO_2 gross production (µM C/day).
NPCO2	NUMBER(4,2)	pCO_2 nett production (µM C/day).
SNPCO2	NUMBER(4,2)	Standard error of pCO_2 nett production (µM C/day).
SRPCO2	NUMBER(4,2)	Standard error of pCO_2 respiration (µM C/day).
ROX	NUMBER(4,2)	Oxygen respiration (µM/day).
SROX	NUMBER(4,2)	Standard error of oxygen respiration (µM/day).
MICROX	NUMBER(4,2)	Oxygen respiration in the microplankton size fraction.
SMICROX	NUMBER(4,2)	Standard error of oxygen respiration in the microplankton.
NANROX	NUMBER(4,2)	Oxygen respiration in the nanoplankton size fraction.
SNANROX	NUMBER(4,2)	Standard error of oxygen respiration in the nanoplankton.
PICROX	NUMBER(4,2)	Oxygen respiration in the picoplankton size fraction.
SPICROX	NUMBER(4,2)	Standard error of oxygen respiration in the picoplankton.

Notes

The experiment reference provides a linkage between the metadata held in table OXYHDR and the individual uptake measurements held in OXYDAT.

The source (position and depth) of the incubated water may be identified through IBTTLE. Note that IBTTLE will not be unique for every record in cases where a common water sample was incubated at several depths.

The fields DEPTH and LIGHT are provided as alternative indicators of the conditions under which the sample was incubated. For in-situ incubations, LIGHT will generally be null and DEPTH represents the actual depth of incubation. For on-deck experiments, LIGHT represents the percentage of ambient light reaching the sample: i.e. the light transmission of the incubation screen. DEPTH is computed from this using either CTD downwelling irradiance or beam attenuance data.

The definitions of micro, nano and picoplankton vary from time to time depending on the filters used in the experiment. The definitions for a given experiment are given in the OXYHDR record.

Note that the units are quoted in terms of uptake per day. This is a loose definition. Strictly speaking, the uptake is quoted over the period of the incubation duration. Normally this is approximately 24 hours but users are advised to check the duration in the appropriate field of OXYHDR.

Table OXYHDR
Field Definitions

EXPREF	CHAR(6)	BODC experiment reference.
TYPE	CHAR(2)	Experiment type code. (OD for on deck experiments, IS for in-situ experiments)
BENCOL	NUMBER(6)	BODC event number of the water collection event.
BEN	NUMBER(6)	BODC event number assigned to the incubation..
SDATE	DATE	Date and time of the start of the incubation.
INCDUR	NUMBER(3,1)	Incubation duration (hours).
COMM	CHAR(30)	Plain language comment field.
DEPINT	NUMBER(4,1)	Depth over which the integrated production was calculated.
INTGOX	NUMBER(6,2)	Integrated gross oxygen production(mmol/m 2/day)
INTNOX	NUMBER(6,2)	Integrated nett oxygen production(mmol/m 2/day)
IGTCO2	NUMBER(6,2)	Integrated gross TCO_2 production(mmol/m 2/day)
INTCO2	NUMBER(6,2)	Integrated nett TCO_2 production(mmol/m 2/day)
MICDEF	CHAR(8)	Microplankton definition.
NANDEF	CHAR(8)	Nanoplankton definition.
PICDEF	CHAR(8)	Picoplankton definition.

Notes

Fields BENCOL and BEN require some explanation as the presence of two BODC event numbers in a single table may at first sight seem confusing.

BENCOL specifies where the water used in the production experiment came from. In some ways it is superfluous because the same information may be derived from the IBTTLE field in OXYDAT. However, it is included to simplify the task of linking integrated production data held in table OXYHDR to the place and time to which they relate.

BEN is a reference given to some production experiments. This invariably relates to *in situ* experiments where a rig has been cast adrift from the ship. On-deck incubations have never been considered as events. The reason for this is more historical than logical: the event entries are drawn up from ship’s logs and whilst a rig being deployed has often (but not always) merited a log entry, the placing of samples in an on-deck incubator has not.

Integrated production data are only included if they were computed and supplied by the data originator. They are not routinely determined by BODC.

Table ORGCODE
Field Definitions

IORGRF	NUMBER(6)	Originator's reference code.
CORGNM	CHAR(20)	Originator's name.
CORGO	CHAR(40)	Originator's organisation.

Table PRINDX

Field Definitions

BEN	NUMBER(6)	BODC event number.
IPROF	NUMBER(6)	BODC profile identifier.
TBEGNS	DATE	Start time of the profile.
MAXDPTH	NUMBER(4,1)	Maximum depth reached (m).
SAMPINT	NUMBER(3,1)	Sampling interval of an instrument (s).
UD	CHAR(1)	Discriminator for upwelling/downwelling (U or D or above-surface downwelling irradiance (S).
RI	CHAR(1)	Discriminator for radiance (flat sensor)/irradiance (hemispherical).
WLMIN	NUMBER(4,1)	Minimum recording wavelength (nm).
WLMAX	NUMBER(4,1)	Maximum recording wavelength (nm).

Table PRPROF

Field Definitions

IPROF	NUMBER(6)	BODC profile identifier.
CYCLE	NUMBER(4)	Profile cycle number.
DEPTH	NUMBER(6,3)	Depth (m).
RAD	NUMBER(6,3)	Radiation intensity ($\mu W/cm^2/sr/nm$)

Notes

IPROF is included to allow for the case of several radiometers being attached to a single frame and lowered as one cast.

The CYCLE field just lists the cycle number of the particular profile in consecutive numbers starting from number one. Note that this relates to a sampling interval which is stored in PRINDX table.

Table ZUCT

Field Definitions

CCTREF	CHAR(4)	Category code.
CCTFUL	CHAR(40)	Category description in plain language.
CILOAD	CHAR(6)	Date of record creation (yymmdd).
TCTMOD	DATE	Record modification time stamp.

Notes

The category codes are designed to group parameters into logical subgroups according to general operational practices. However, there will inevitably be parameters that could be fitted into more than one category depending upon one’s point of view. This should be borne in mind when searching the dictionary. Always check out all possible categories.

Table ZUNT
Field Definitions

CPUREF	CHAR(4)	Unit code.
CPUABB	CHAR(10)	Abbreviated unit description.
CPUFUL	CHAR(40)	Full unit description.
CILOAD	CHAR(6)	Date record was created (yymmdd).
TPUMOD	DATE	Last modification time stamp.

Table ZUPM
Field Definitions

CPMCAT	CHAR(4)	Category code.
CPMREF	CHAR(4)	4-byte code for the parameter name.
CPMABB	CHAR(20)	Abbreviated parameter name.
CPMFUL	CHAR(80)	Full parameter name.
CPMUNT	CHAR(4)	Parameter storage unit code.
FABSNT	NUMBER	Absent data value.
FPMINM	NUMBER	Minimum value for parameter.
FPMAXM	NUMBER	Maximum value for parameter.
CINVER	CHAR(1)	Plot inversion flag.
CILOAD	CHAR(6)	Date of record creation (yymmdd).
TPMMOD	DATE	Date and time of last modification.

Notes
Most of the fields in this table are of more interest to BODC personnel than to database users. The exceptions are CPMCAT, CPMREF, CPMFUL and CPMUNT.

The category code (CPMCAT) provides the linkage to table ZUCT and hence identifies which generalised parameter descriptions belong to which category. CPMFUL contains the parameter description in plain language and provides the hook by which users can recognise just what is meant by a particular code.

The field CPMUNT specifies the units in which the parameter is stored in the database. This is present as a code (to prevent problems arising from differing descriptions being given to the same unit e.g. degrees, deg. and the like) which may be translated using table ZUNT.

Table ZUSG
Field Definitions

CPMREF	CHAR(4)	Parameter name code (bytes 1-4)
CSGREF	CHAR(2)	Parameter subgroup code (bytes 5-6).
CDSREF	CHAR(2)	Parameter discriminator code (bytes 7-8).
CPMUSG	CHAR(8)	Full 8-byte parameter code.
IPMBEF	NUMBER(1)	Number of digits before the decimal point.
IPMAFT	NUMBER(1)	Number of digits after the decimal point.
CSGABB	CHAR(20)	Abbreviated parameter code description.
CSGFUL	CHAR(100)	Full parameter code description.
CSGMTH	CHAR(100)	Methodology description.
ISGREF	NUMBER(8)	Narrative document reference.
CILOAD	CHAR(6)	Date record was created (yymmdd).
TSGMOD	DATE	Record modification time stamp.

Notes

The complete parameter code (CPMUSG) is constructed by concatenation of the parameter name, parameter subgroup and parameter discriminator codes.

The fields IPMBEF and IPMAFT are included to allow software to format data sensibly. Note that the data covered by the parameter codes are stored to a precision of some 16 decimal places. IPMAFT indicates how many of these have significance.

The meaning of a given code is specified in plain language by the fields CSGFUL and CSGMTH. These fields are designed to give a user-friendly reference to the full parameter code. If they don't, please let us know. All the details which make the parameter unique (including filtration details where appropriate) are included.

The ISGREF field allows a linkage point for data documentation. It is designed to allow general methodology description documents to be linked to a parameter code.

This on-line documentation is not currently implemented and the field is set null.

Database Linkage Definitions

The tables in this section of the document show the linkages that exist between the database tables. The linkages chains run along the rows of the table and always start with table EVENT. The type of linkage is shown by bolding the text. A linkage from normal text to bold text is a 'one to many' relationship. Links from normal text to normal text or bold text to bold text are 'one to one' relationships.

CTD Data

EVENT	CTDINDX	CTDCAL	CTDTYP	BINCTD
BEN	BEN	BEN	BEN	**BEN**
	TYPE		TYPE	

Water and Air Sample Data (Fully Normalised)

EVENT	BOTTLE	BOTDATA	ZUSG	ORGCODE
BEN	**BEN**			
	IBTTLE	**IBTTLE**		
		IORGRF		IORGRF
		CPCODE	CPMUSG	

TCO2/ O2 Production Data

EVENT	BOTTLE	OXYDAT	OXYHDR
BEN	**BEN**		
	IBTTLE	**IBTTLE**	
		EXPREF	EXPREF
BEN			BENCOL

Irradiance Profiles

EVENT	PRINDX	PRPROF
BEN	**BEN**	
	IPROF	**IPROF**

Parameter Dictionary

ZUCT	ZUPM	ZUNT	ZUSG
CCTREF	**CPMCAT**		
	CPMUNT	CPMUNT	
	CPMREF		**CPMREF**